

Multiple and Dissociable Effects of Sensory History on Working-Memory Performance

Jasper E. Hajonides,^{1,2}  Freek van Ede,³ Mark G. Stokes,^{1,2}  Anna C. Nobre,^{1,2} and Nicholas E. Myers⁴

¹Oxford Centre for Human Brain Activity, Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, OX3 7JX, United Kingdom,

²Department of Experimental Psychology, University of Oxford, Oxford, OX2 6GG, United Kingdom, ³Department of Applied and Experimental Psychology, Vrije Universiteit Amsterdam, 1081 BT, Amsterdam, Netherlands, and ⁴School of Psychology, University of Nottingham, Nottingham, NG7 2RD, United Kingdom

Behavioral reports of sensory information are biased by stimulus history. The nature and direction of such serial-dependence biases can differ between experimental settings; both attractive and repulsive biases toward previous stimuli have been observed. How and when these biases arise in the human brain remains largely unexplored. They could occur either via a change in sensory processing itself and/or during postperceptual processes such as maintenance or decision-making. To address this, we tested 20 participants (11 female) and analyzed behavioral and magnetoencephalographic (MEG) data from a working-memory task in which participants were sequentially presented with two randomly oriented gratings, one of which was cued for recall at the end of the trial. Behavioral responses showed evidence for two distinct biases: (1) a within-trial repulsive bias away from the previously encoded orientation on the same trial, and (2) a between-trial attractive bias toward the task-relevant orientation on the previous trial. Multivariate classification of stimulus orientation revealed that neural representations during stimulus encoding were biased away from the previous grating orientation, regardless of whether we considered the within-trial or between-trial prior orientation, despite opposite effects on behavior. These results suggest that repulsive biases occur at the level of sensory processing and can be overridden at postperceptual stages to result in attractive biases in behavior.

Key words: biases; machine learning; MEG; neural representations; serial dependence; working memory

Significance Statement

Recent experience biases behavioral reports of sensory information, possibly capitalizing on the temporal regularity in our environment. It is still unclear at what stage of stimulus processing such serial biases arise. Here, we recorded behavior and neurophysiological [magnetoencephalographic (MEG)] data to test whether neural activity patterns during early sensory processing show the same biases seen in participants' reports. In a working-memory task that produced multiple biases in behavior, responses were biased toward previous targets, but away from more recent stimuli. Neural activity patterns were uniformly biased away from all previously relevant items. Our results contradict proposals that all serial biases arise at an early sensory processing stage. Instead, neural activity exhibited mostly adaptation-like responses to recent stimuli.

Received June 19, 2022; revised Nov. 30, 2022; accepted Jan. 19, 2023.

Author contributions: M.G.S., A.C.N., and N.E.M. designed research; M.G.S. and N.E.M. performed research; J.E.H. contributed unpublished reagents/analytic tools; J.E.H. and N.E.M. analyzed data; J.E.H., F.v.E., M.G.S., A.C.N., and N.E.M. wrote the first draft of the paper; J.E.H., F.v.E., A.C.N., and N.E.M. edited the paper; J.E.H., F.v.E., and N.E.M. wrote the paper.

This research was supported by an Economic and Social Research Council Grand Union studentship and the Scatcherd European Scholarship (J.E.H.), the European Research Council Starting Grant MEMTICIPATION 850636 (to F.v.E.), the James S. McDonnell Foundation Scholar Award 220020405 and the ESRC Grant ES/S015477/1 (to M.G.S.), the James S. McDonnell Foundation Understanding Human Cognition Collaborative Award Number 220020448 and the Wellcome Trust Senior Investigator Award 104571/Z/14/Z (to A.C.N.), and the Wellcome Trust Award 201409/Z/16/Z and support from University College Oxford (N.E.M.). The work was enabled by the National Institute for Health Research Oxford Health Biomedical Research Centre and the Wellcome Centre for Integrative Neuroimaging supported by core funding from the Wellcome Trust Grant 203139/Z/16/Z. For the purpose of Open Access, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

The authors declare no competing financial interests.

Correspondence should be addressed to Nicholas E. Myers at nicholas.myers@nottingham.ac.uk.

<https://doi.org/10.1523/JNEUROSCI.1200-22.2023>

Copyright © 2023 Hajonides et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Introduction

Stimulus history modulates performance guided by sensory input. Reliance on temporal correlations is deeply engrained in the visual system (Simoncelli and Olshausen, 2001) and can be beneficial to guide perception within a world that is largely stable over short time scales (Dong and Atick, 1995). Past sensory evidence can be an effective prior for extracting signals from the noisy sensory stream and to help maintain a stable representation to bridge blinks, eye movements, or visual occlusions.

Recent studies have also revealed robust short-term influences of prior stimulation in delayed-response and working-memory tasks (Huang and Sekuler, 2010; Cicchini et al., 2014, 2018; Fischer and Whitney, 2014; Bae and Luck, 2017; Fritsche et al., 2017; Czoschke et al., 2019). More specifically, remembered items are often reported as more similar to the task-relevant stimulus in a previous trial. This attractive bias, sometimes referred to as the serial-dependency bias, is suggested to increase

temporal stability by acting as a prior (Cicchini et al., 2014; Fischer and Whitney, 2014; Kiyonaga et al., 2017; Fischer et al., 2020). Conversely, features of items sequentially presented within trials are often judged as more dissimilar to each other (Born and Tootell, 1992; Störmer and Alvarez, 2014; Fritzsche et al., 2017). This repulsive bias could result from efficient coding of temporally autocorrelated signals (Cicchini et al., 2018). The amplification of subtle differences between stimuli could additionally serve to optimize perceptual decision-making (Kiyonaga et al., 2017; Cicchini et al., 2018; van Bergen and Jehee, 2019). Despite their opposite directions, attractive and repulsive performance biases have been shown to jointly influence task processing (Fritzsche et al., 2017, 2020; Czoschke et al., 2019), albeit at different time scales (Gekas et al., 2019; Sheehan and Serences, 2022).

While the current literature has largely focused on behavior, relatively less work has examined whether these opposing biases share a neural mechanism, and in particular whether both biases may arise at early or late sensory processing stages. A number of studies have shown evidence for between-trial biases arising in early visual cortex (St. John-Saaltink et al., 2016; Sheehan and Serences, 2022), but with insufficient temporal resolution to isolate early processing stages, or they have measured activity only in frontal cortex but not sensory cortex (Papadimitriou et al., 2017).

The current study combined behavioral reports in a working-memory task and continuous magnetoencephalographic (MEG) recordings to assess within-trial and between-trial biases induced by sensory history with high temporal resolution. We devised a task in which participants were sequentially presented with two orientations and used a continuous, precision response to reproduce one of them at the end of a trial. Additionally, we included a cue that indicated which of the two orientations had to be reported in the trial to test for possible differential susceptibility of biases to the task-relevance of stimuli (as per Bae and Luck, 2020; Fischer et al., 2020). We compared systematic biases induced by the first orientation on the second (within-trial bias) as well as biases from the previous trial on the current trial (between-trial bias). Critically, we tested whether we could find neural signatures for such biases using a decoding approach that leveraged the high temporal resolution of MEG. We set out to decode the presented orientation and expected to find the within-trial and between-trial biases in the neural data that mirrored the behavioral biases.

Previewing the results, we confirmed that behavioral responses were systematically pushed away from previous orientations on the same trial and pulled toward the orientations recalled on the previous trial. Mirroring the behavioral repulsion results, the sensory neural representation of stimulus orientation was shifted away from the preceding orientation presented on a given trial. However, we found no neural evidence for an attractive between-trial sensory bias during orientation encoding, despite such a bias observed in behavior. Instead, we always observed a repulsive neural bias, regardless whether we considered orientations from the same or previous trial as a source of the bias.

Materials and Methods

Participants

Twenty healthy volunteers with normal or corrected-to-normal vision participated in the study. All participants were between 20 and 36 years old (mean 25.4 years old; 11 females). Before taking part in the study, volunteers provided their informed consent according to the procedures approved by the Central University Research Ethics Committee of the

University of Oxford. Participants received £15 per hour compensation for taking part in this study.

Experimental set-up

Participants sat in the MEG scanner, which was situated in a dimly lit, sound-proof, and magnetically shielded room. A projection screen was placed at a viewing distance of 90 cm. Visual stimuli were projected at the back of the screen at a spatial resolution of 1024×768 pixels using a refresh rate of 60 Hz using a Panasonic DLP projector (PT-D7700E).

The task was programmed and presented using MATLAB (MathWorks) in conjunction with the Psychophysics Toolbox (Brainard, 1997). Participants indicated their responses on an optic-fiber response box.

Task

Participants performed a precision working-memory task in which they reproduced the orientation of one of two grating stimuli presented sequentially with independent orientations (Fig. 1). Simultaneously with the presentation of the second grating, participants were cued as to which grating orientation to report when probed at the end of the trial. In half of the trials, only the first or second grating was presented and used for reporting.

Each trial started with a central fixation dot (0.2° visual angle) on screen for 800 ms with a gray background (RGB: 127, 127, 127; Fig. 1). Subsequently, a sinusoidal Gabor stimulus at a random angle was centrally presented on the visual display for 200 ms (diameter of 6° visual angle, two cycles per degree of visual angle, 50% contrast, tapered by a Gaussian envelope with a 1.5° SD). On trials in which the first grating was not presented the fixation dot changed color from black to gray (RGB: 192, 192, 192) to signal the omission. After a delay of 1700–1900 ms, the second grating was presented. This second grating had the same properties as the first grating except for the angle, which was randomly drawn independently from that of the first grating orientation. At the center of the grating, the fixation dot changed color [orange: (255, 161, 0); cyan: (0, 236, 255)]. The color signaled if the first orientation (report first trial) or the second orientation (report second trial) would be probed. Cue-color contingencies of the experiment were counterbalanced across participants and changed halfway through the experiment. After the color contingencies switched, participants practiced the new contingencies over the course of one block before continuing with the second half of the session. The color cue was always valid in indicating which grating orientation was relevant for the reporting stage at the end of the trial. On trials in which the second grating was absent, the cue would therefore always indicate the first orientation.

After another delay of 1700–1900 ms, a probe grating appeared. Participants had to adjust its orientation to match the cued grating in memory. Adjustments were made by pressing buttons with the right hand to rotate the grating either clockwise (CW; middle finger) or anti-clockwise (index finger). Responses were confirmed by pressing a button with the left index finger. A 200-ms fixation period followed, after which participants were presented with 50 ms of feedback in the form of a grating indicating the correct orientation.

In total, participants completed 400 trials. In half of the trials, two items were presented in 200 trials (100 trials with report first cue and 100 with report second cue). Only one item was presented in the remaining 200 trials (100 with only grating 1 and 100 only grating 2). The resulting factorial design (relevant grating, number of gratings presented) included four conditions: report first with two items presented, report first with one item presented, report second with two items presented, and report second with one item presented. Trial types were randomly mixed and presented in blocks of 50 trials, with each block lasting ~ 10 min.

Behavioral analysis

Response error was quantified by calculating the circular distance between the recalled orientation and the cued orientation. All responses were mapped onto a -90° to 90° space to compute circular error. In all between-trial analyses, we excluded the first trial of each block.

Mixture modeling

We fit a classical mixture model frequently used in the working memory literature to investigate the relative rate of responses to targets, guesses,

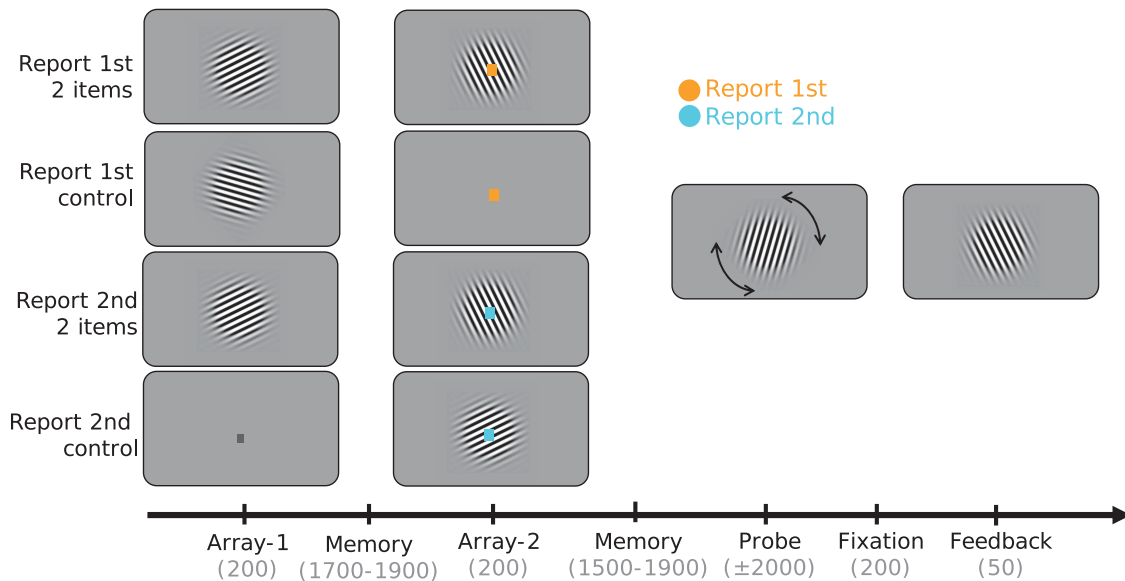


Figure 1. Experimental task design. Participants were presented with two arrays that contained an oriented bar grating or a place holder fixation dot. An oriented grating was presented either in the first (25% of trials), second (25%), or both (50%) arrays. A color cue presented on array 2 indicated the relevant stimulus for recall. The cue size in array 2 is enlarged for visualization purposes. The color cued either the first array or the second array. On trials with only one orientation presented, the cue was redundant but still presented. Mappings between color and cue meaning were counterbalanced. After matching the orientation in the probe display to the orientation in memory, participants were presented with feedback in the form of a grating with the correct orientation. Timings in parentheses are in milliseconds.

and erroneous responses to the wrong target (“swap errors”; Bays et al., 2009; Schneegans and Bays, 2017). The model was fit separately for each condition (report first in two-item trials; report first in one-item control; report second in two-item trials; report second in one-item control). The mixture model estimated the precision of the von Mises distribution, target response rate, guess rate, and swap rate to the item that was presented on the same trial but not cued (only two-item trials). After fitting the data to the entire dataset, the model provided single-trial weights for target response rate, guess rate, and swap rate.

Performance bias calculation

We calculated the performance bias (the signed circular difference between the response and the target orientation) as a function of the circular difference between the target orientation and the orientation that induced the bias (the target orientation in the preceding trial or the task-irrelevant orientation in the same trial). We did not use response orientation to avoid confounding serial bias with the oblique bias, which could artificially inflate the bias estimate (De Gardelle et al., 2010; Tomassini et al., 2010). We computed the difference between the target orientation and inducer through subtraction, with all angular differences mapped between -90° and 90° . These distances were binned into 64 equally sized, overlapping bins, with each bin containing 25% of trials. We computed the average signed error (bias) across all trials within each bin. Edge artefacts were avoided by wrapping the angular differences around, to ensure the average was computed over a constant number of trials and centered on the mid-point of the bin. Subsequently, we sign-flipped the performance bias in bins with a negative distance and averaged over negative (-90° to 0°) and positive distances (0° to 90°) between target orientation and inducing orientation, resulting in 32 bins. The summed bias over absolute distances (0° to 90°) calculated over the 32 bins for each condition and participant served as a measure of bias.

MEG acquisition

Participants were seated in the MEG scanner after being instructed about the task specifics. They completed one practice block while seated in the scanner before MEG recording onset. Participants were instructed to maintain their gaze at the central fixation dot and to minimize blinking throughout the trial.

Neuromagnetic data were acquired using a whole-head VectorView system including 204 planar gradiometers and 102 magnetometers (Elekta Neuromag Oy) in a magnetically shielded room. Throughout the

experiment, participants’ head position was monitored continuously using index coils placed at four points on the head. Magnetic field strength was sampled at a rate of 1000 Hz and bandpass filtered online between 0.03 and 300 Hz. In addition, vertical and horizontal electro-oculograms were measured using electrodes placed above, below, and adjacent to the eyes. Eye movements were monitored using an EyeLink 1000 (SR Research) eye tracker at a frequency of 1000 Hz.

MEG data preprocessing

The data were preprocessed offline using Fieldtrip (Oostenveld et al., 2011), OHBA software library (OSL) drawing on SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>), and Elekta software. Before any preprocessing, the MEG data were visually inspected to remove and interpolate any sensors that displayed excessive levels of noise and were subsequently de-noised and motion corrected using Maxfilter Signal Space Separation (Taulu et al., 2004) before removing independent components related to cardiac and eye-blink artefacts. Data were epoched around the first grating and second grating (from 400 ms before grating onset to 900 ms after onset) and downsampled to 200 Hz. Trials with high variance in either gradiometers or magnetometers were identified and excluded using a generalized ESD (extreme studentized deviate; Rosner, 1983) test at a 0.05 significance threshold.

For between-trial bias analyses, we trained the classifier on all stimulus presentations from all nonrejected trials. Since we excluded trials with high variance ($7.49 \pm 3.85\%$, mean \pm SD, corresponding to 45 ± 23 stimulus presentations) from further analysis, the classifier was trained on the remaining 555 ± 23 stimulus presentations. Before calculating biases, we removed the first trial of each block (2%), and trials with an absolute angular difference with the target orientation on the previous trial of $>60^\circ$ ($\sim 33\%$), resulting in 370 ± 21 trials for this analysis per participant.

For within-trial bias analyses, we trained the classifier on all trials where stimulus 2 was presented except for those removed for high variance, leaving 278 ± 11 trials in the analysis. For the bias calculation, among these trials we selected those on which stimulus 1 had also been presented, and with an absolute angular difference of $>10^\circ$ and $<50^\circ$ between stimuli 1 and 2. Subsequently, the bias was calculated separately for trials with protect and update cues, resulting in 62 ± 5 and 63 ± 5 trials, respectively.

Linear discriminant analysis classification

Data were further preprocessed. Magnitudes of magnetometers were approximately matched to gradiometers by multiplication (factor 20) and subjected to spatiotemporal decoding (code available at <https://pypi.org/project/temp-dec/>; as described previously, Wolff et al., 2017, 2020; Hajonides et al., 2021). Data from all 306 MEG sensors across a sliding window of 30 time points (150 ms) were concatenated into a 9180-dimensional vector. Dimensionality was reduced using principal component analysis, computed separately for each time point, maintaining 90% of the variance (between 250 and 600 ms, this was around 209 ± 39 components per participant, mean \pm SD). This served to de-noise the data, increase the unique variance encoded by each dimension, and enable the classifier to compute covariance matrices more effectively. Prestimulus baselining was not applied to maintain stable information from previously presented stimuli.

To train a linear discriminant analysis (LDA) classifier, the data were split into training and testing sets using 10-fold stratified cross-validation. Grating angles were binned into equally spaced orientation bins, creating 10 distinct classes (0–18°, 18–36°, 36–54°, 54–72°, 72–90°, 90–108°, 108–126°, 126–144°, 144–162°, 162–180°). To train a LDA classifier, the data were split into training and testing sets using 10-fold stratified cross-validation. Based on the training set, the LDA classifier projects the data into a low-dimensional space (of nine dimensions; number of classes minus 1) that maximally separates the data from the 10 classes. Data from the test set were then projected into the same 9-dimensional space. We obtained 10 LDA distances for each trial in the test set by calculating its Euclidean distance from each training set class mean in the low-dimensional space. These distances allowed us to estimate the likelihood that any given test trial corresponded to each of the ten classes. This was repeated for each cross-validation fold and each time point. In stimulus decoding analyses, the presented angle was used for classification. In cross-decoding analyses, LDA classifiers were trained on orientation bins of one event (e.g., presented grating) but classifier evidence aligned around bins of another orientation (e.g., target orientation on the previous trial). The resulting representational similarity curves were convolved with a cosine.

To test which sensors most significantly contributed to the classifier likelihoods observed in our multivariate methods, we also ran a searchlight decoding analysis (Kriegeskorte et al., 2006). In this analysis, we iteratively considered a small group of sensors and were thereby able to map the approximate locus of the observed effect. More specifically, we selected data from each sensor plus its 47 most closely adjacent neighbors (magnetometers and gradiometers included) and ran the same classification analysis as described above.

Calculation of the neural asymmetry score as a measure of neural bias

For within-trial biases, we assessed processing of the second grating and only considered two-item trials. The classifier was trained on all presentations of the second grating and bin likelihoods were generated for each trial. For between-trial analyses, we analyzed orientation processing of both the first and second grating in the current trial. For this reason, we trained the classifier on all trials and generated bin predictions for all trials.

Subsequently, based on the results from the performance-bias analyses, we selected trials in which the angular distance between the inducer and the grating orientation on the display led to a significant behavioral bias at the group level. In the case of the within-trial repulsive bias, the inducer was the orientation of the first grating on the same trial. For the between-trial analyses, the inducer was the target orientation reported on the previous trial (except for control analyses, where the unreported orientation was used as the target). As a dependent variable, we considered likelihood estimations for each orientation bin, where we expect the highest likelihood for the angular bin that has zero offset to the presented orientation and decreasing likelihoods for bins with larger angular distances to the presented orientation. We separately assessed likelihood estimations for trials in which the inducer orientation was clockwise (CW) versus counterclockwise (CCW) with respect to the current orientation. For both CW and CCW trials, we separately averaged the evidence from the orientation bins CW (–72° to –18°) and evidence

Table 1. Descriptive statistics of error for all experimental conditions

	Report 1st trials	Report 2nd trials
1 item	11.91° ($\pm 4.61^\circ$)	8.96° ($\pm 2.52^\circ$)
2 items	14.69° ($\pm 5.38^\circ$)	9.79° ($\pm 2.89^\circ$)

Mean absolute error (\pm SD). All units are in degrees.

from the CCW bins (18° to 72°). Asymmetry scores were computed by obtaining the difference between the two groups of angular bins (CW minus CCW). Finally, we calculated an overall neural bias score by subtracting asymmetry scores on trials with CW versus CCW inducers. Attractive neural biases resulted in a positive score (i.e., trials with CW angular distances resulted in more CW evidence, CCW angular distances resulted in more CCW evidence), whereas repulsive neural biases resulted in a negative score (i.e., CW angular distances resulted in less CW evidence than CCW trials, and vice versa).

Statistical testing

Statistical tests were computed using both JASP (JASP Team, 2020) and Scipy (Virtanen et al., 2020).

We tested the time series of cosine-convolved classifier evidence against zero using a cluster-based permutation test, which addresses the multiple comparison problem (using MNE; Gramfort et al., 2013). We ran 100,000 iterations. The clusters with groups of time points significantly different from zero are indicated in the relevant figures using horizontal lines. Cluster-based permutation testing was also applied to performance bias across angular distance between the presented orientation and the inducer orientation.

To test the significance of our bias analyses, we generated a shuffling distribution after the decoding stage. When trials were sorted based on the relative orientation of the previous trial/stimulus, we randomly sign-flipped this angular distance and re-calculated the bias. We computed the bias for all participants and averaged this score. This process was repeated 10,000 times, and the resulting distribution was z-scored. The same z score transformation was applied to the observed bias score when no random sign-flipping was applied. This z score could then be used to get the (two-tailed) *p*-value of the original effect relative to the shuffling distribution (in all reported time averages, time points between 250 and 600 ms were used).

All tests were two-sided unless stated otherwise.

Data availability

Data and analysis scripts can be accessed via <https://osf.io/98ujm/>.

Results

Error rates

Participants were accurate in reproducing the target orientation (mean response error $11.73 \pm 0.70^\circ$ SEM; mean SD $17.61 \pm 1.07^\circ$ SEM; see Table 1 for condition-wise performance). A two-by-two repeated-measures ANOVA on response error showed main effects of cue type ($F_{(1,19)} = 16.49$, $p < 0.001$, $\eta^2 = 0.374$) and number of stimuli presented ($F_{(1,19)} = 29.78$, $p < 0.001$, $\eta^2 = 0.075$). Cue type was significant for both one-item or two-item trials, with absolute error higher on report first than on report second trials for both two-item trials ($t_{(19)} = 3.972$, $p < 0.001$, $d = 0.888$; see Table 1) and one-item trials ($t_{(19)} = 3.948$, Bonferroni-corrected $p < 0.001$, $d = 0.883$). By contrast, number of items presented primarily affected the report first conditions. Error was higher on report first two-item than on report first one-item trials ($t_{(19)} = 5.665$, $p < 0.001$, $d = 1.267$) but did not significantly differ between report second two-item and report second one-item trials ($t_{(19)} = 1.885$, $p = 0.075$, $d = 0.421$), leading to a significant interaction between the two factors ($F_{(1,19)} = 10.90$, $p = 0.004$, $\eta^2 = 0.026$). Analyses using mixture modeling (Bays et al., 2009) confirmed that errors originating from responses to the noncued grating orientation

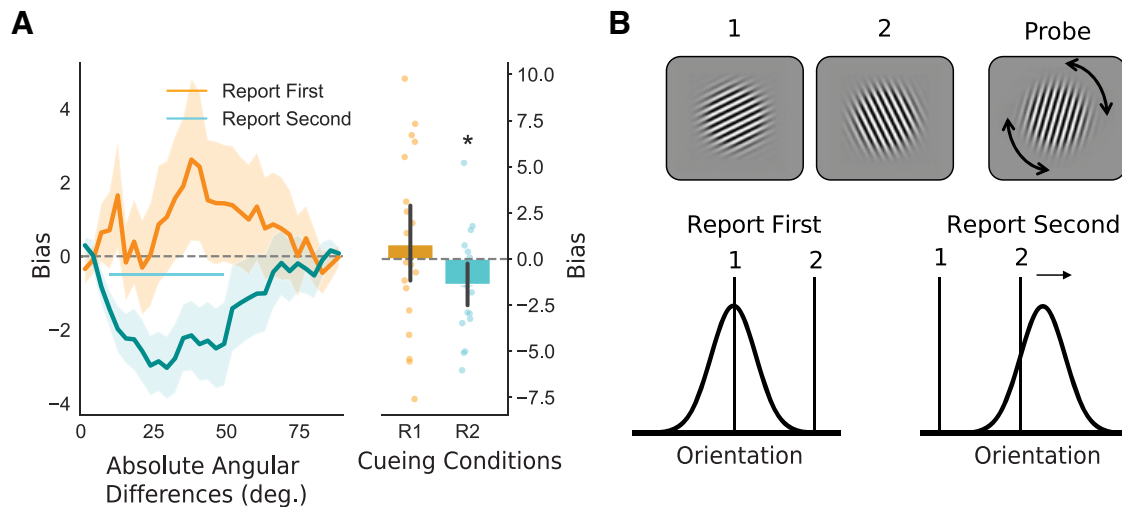


Figure 2. Within-trial repulsive performance biases. **A**, Within-trial bias as a function of the absolute angular distance between the two presented orientations. Trials were split for report first and report second cues on two-item trials. Shading indicates SEM. The horizontal line indicates a cluster of significant performance biases for a range of angular difference. The right panel shows the sum of bias across angular distances per subject. Error bars show 95% confidence intervals. R1 = report first; R2 = report second. **B**, Schematic of the within-trial performance bias. A repulsive performance bias is observed for report second trials only. * indicates $p < 0.05$.

were rare (swap rate of 0.033 ± 0.01 on two-item trials; see also Huang, 2020).

Repulsive performance biases within trials

We analyzed within-trial biases in behavioral performance by assessing whether the reported orientation was systematically reported as closer to or further away from the nontarget orientation on the same trial (see Materials and Methods). We restricted the analyses to two-item trials. Figure 2A shows the performance bias for all absolute angular distances between the first and second grating orientation for report first and report second trials. In trials with report first cues, there was no significant bias toward or away from the interfering second grating orientation that was not relevant to the task at hand ($t_{(19)} = 0.74$, $p = 0.467$). In contrast, trials with report second cues revealed significant biases away from the initially encoded first grating orientation ($t_{(19)} = -2.33$, $p = 0.031$; illustrated in Fig. 2B). The repulsive bias in report second trials was confirmed using a cluster-based permutation test, showing a significant cluster ($p = 0.012$) when the angular distance between the two orientations was between 10° and 49° (Fig. 2A).

Attractive performance bias between trials

We next evaluated the between-trial bias on responses in the current trial toward the orientation that was cued on the previous trial (Fig. 3). We assessed the performance bias as a function of angular differences between the target grating on the current and on the previous trial. The analysis also considered the position of the target grating in the current trial (first or second) and the number of items on the current trial (one-item or two-item). For consistency, we term all trials where participants report the first grating orientation report first trials and trials where participants report the second orientation report second trials, regardless of the number of gratings presented. Again, we calculated the sum of the bias across angular distances between targets in the current and previous trial (Fig. 3A,B). In contrast to the repulsive bias described in the previous section, we found that all conditions showed an attractive performance bias (all $p < 0.05$ in two-sided statistical tests). The attractive serial bias was most pronounced for small to intermediate angular distances between the inducer

and current orientation (0 – 60°). A repeated-measures ANOVA on the sum of biases across angular distances indicated an effect of cue type, with larger biases occurring in report first trials ($F_{(1,19)} = 5.706$, $p = 0.027$, $\eta^2 = 0.172$), but not of the number of gratings presented in a trial ($F_{(1,19)} = 0.980$, $p = 0.335$, $\eta^2 = 0.007$). The two factors did not interact ($F_{(1,19)} = 0.377$, $p = 0.547$, $\eta^2 = 0.002$). This shows that the bias was stronger when recalling the first item, which was encoded closer in time to the previous trial.

Task dependence of attractive performance bias between trials

Next, we repeated the same between-trial analyses but investigated the role of task relevance and cue type in the previous trial. This allowed us to test for and compare behavioral biases elicited by the probed (and reported) orientation and by the unreported orientation in the previous trial. We also tested whether the cue type on the previous trial affected bias in the current trial. We only looked at trials in which two orientations were presented on the previous trial.

A repeated-measures ANOVA on the sum of angular distances of the performance bias for task relevance and cue type confirmed an effect of task relevance ($F_{(1,19)} = 14.684$, $p = 0.001$; Fig. 3D) but showed no effect of cue type ($F_{(1,19)} = 1.423$, $p = 0.248$) or interaction ($F_{(1,19)} = 1.633$, $p = 0.216$). Task-relevant orientations in trials of either cue type resulted in a significant bias (previous trial was report first and contained two items: $t_{(19)} = 3.524$, $p = 0.002$; previous trial was report second and contained two items: $t_{(19)} = 4.476$, $p < 0.001$). Unprobed orientations did not lead to a significant bias (both $p > 0.2$). No reliable difference was observed between the strength of the bias between report second two-items or report first two-items conditions in previous trials ($t_{(19)} = 1.691$, $p = 0.107$). Following up, we assessed the performance bias as a function of the angular distance between the current target orientation and previously presented orientations. Cluster-based permutation testing showed an attractive bias toward the task-relevant orientation on the previous trial when a report first cue ($p = 0.001$; 7 – 58° ; Fig. 3C) or a report second cue ($p < 0.001$; 4 – 77°) was presented in the previous trial. For unreported orientations, no bias was observed (no candidate

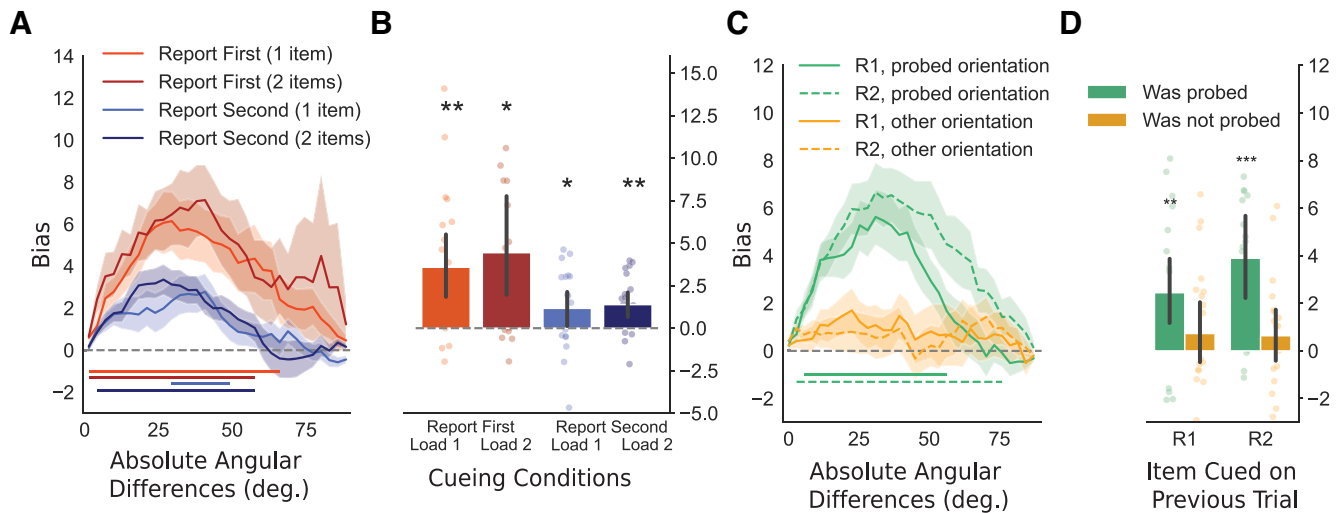


Figure 3. Attractive behavioral biases toward the relevant target orientation in the previous trial. **A**, Line plots for the attractive between-trial bias as a function of the absolute angular distance between the target orientation on the previous trial and the presented grating. Shading indicates the SEM. Horizontal lines indicate clusters of significant performance biases for a range of angular differences. **B**, Bar plots indicate the average sum of biases across angular distances per subject and per condition as seen in panel **A**. **C**, Line plots show the attractive between-trial bias as a function of the absolute angular distance between the probed angle on the previous trial (green) or the presented but not probed angle in the previous trial (orange). The horizontal lines indicate cluster-corrected angular distances for which the bias is significant. R1 = report first orientation on the previous trial; R2 = report second orientation on the previous trial. **D**, Sum of the response data shown in panel **C**. Positive values indicate an attractive bias. All error bars indicate 95% confidence intervals. * $p < 0.05$, ** $p < 0.01$.

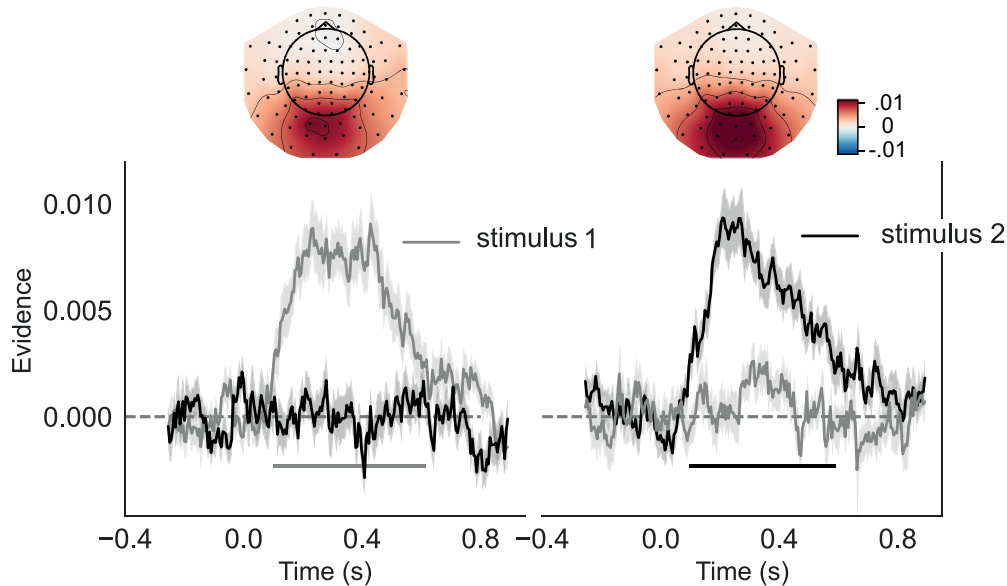


Figure 4. Orientation decoding of presented grating. Training an independent classifier to decode the first or second grating orientation in either the first-time or second-time interval, left and right side, respectively. The MEG topographies show which sensors most strongly contribute to the overall evidence. The left topography shows decoding of the first orientation in the first interval between 250 and 600 ms. The right topography shows decoding of the second orientation at the same latency. Error bars indicate the SEM. Horizontal lines indicate periods of significant grating orientation decoding after cluster-based permutation tests against zero, $p = 0.05$.

clusters for report first or report second cue). Together, this pattern of results is showing an attractive between-trial bias, but only with regard to items that were relevant in the previous trial.

Neural classification of presented orientations

For our classification analysis, grating orientations were binned into 10 equally spaced bins. We applied linear discriminant analysis (LDA) on spatial and temporal features from all 306 MEG sensors ranging from 400 ms before stimulus onset up to 900 ms poststimulus onset (see Materials and Methods). If orientation information was present, LDA likelihood estimations gave rise to representational similarity curves centered on the presented

orientation that could be convolved with a cosine function to result in a single evidence estimation per time point. LDA classification reflected significant evidence for the presented orientation after the visual onset of the grating (Fig. 4). This revealed significant decoding of both the first grating orientation (100–615 ms; $p < 0.001$) and the second grating orientation (95–590 ms; $p < 0.001$). A nonsignificant trend was observed for the evidence of the first grating orientation following the presentation of the second grating (250–600 ms; $t_{(19)} = 1.971$; $p = 0.063$; see also cross-decoding analyses below). There was no difference in classifier evidence after a report first or report second cue (250–600 ms; $t_{(19)} = 0.798$; $p = 0.435$).

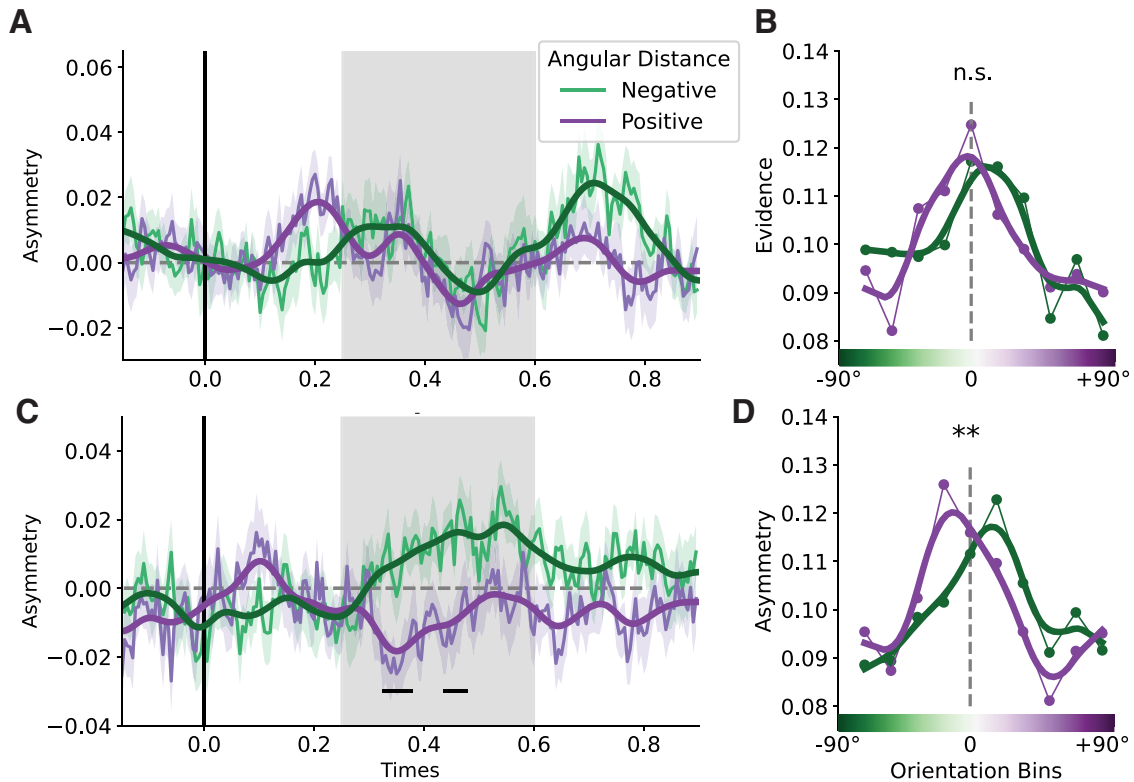


Figure 5. Shift in representational similarity curve relative to previous grating orientation. **A**, Asymmetry scores following the presentation of the second grating on report first two-item trials. The first grating orientation had a positive (purple) or negative (green) angular distance relative to the current grating orientation. The shaded area indicates the time interval (250–600 ms) used for statistical analysis and to generate the representational similarity curves in the right panel. Smoothing (thicker lines in darker colors) was applied for visualization purposes. **B**, Representational similarity curves with evidence for the presented grating orientation for trials with positive or negative angular distances relative to the first rating. For visualization purposes, data points were interpolated (from 10–50 data points) and fitted using a Savitzky–Golay filter (with a window length of nine data points and polynomial of order 1). **C**, Asymmetry scores for report second two-item trials, where participants encoded the grating orientation on screen into working memory with a neural bias away from the grating orientation that was presented earlier on the same trial. Shading indicates SEM. **D**, Data from the gray-shaded region in panel **C**. Horizontal lines indicate periods of significant bias on unsmoothed data after cluster-based permutation tests against $p = 0.05$. * $p < 0.05$, n.s. $p \geq 0.05$.

Within-trial classification bias away from previous stimulus

Representational similarity curves were used to estimate the direction and magnitude of neural biases in orientation representation. By training the classifier on all stimuli, orientation history biases should cancel out, allowing testing on trials with specific prior orientations (clockwise vs counterclockwise to current stimulus) to reveal any neural biases. To investigate how information from the second grating was modulated by information from the first grating, we separately assessed trials with report first and report second cues. Doing so, we evaluated the classification evidence in the MEG data in the epoch following the presentation of the second grating. For these analyses, we again only selected trials with both gratings were presented. We separated trials in which the first grating was clockwise versus counterclockwise relative to the second grating (angular distance of 10–50°, based on behavioral results, see Fig. 2A, B). We considered the average of time points between 250–600 ms for all future analyses, since in this time window stimulus orientation could be decoded with reliable accuracy in this time window (see Fig. 4; see also Fig. 5, gray-shaded area). Echoing the performance biases, no significant bias occurred on report first trials ($Z = -0.645$; $p = 0.516$; Fig. 5A,B). However, on report second trials, we observed a repulsive effect, away from the previously encoded grating orientation ($Z = -2.743$; $p = 0.006$; Fig. 5C,D). There was no correlation, across participants, between the magnitude of the bias in the behavioral

and neural data on report first trials ($r = 0.010$; $p = 0.968$) or report second trials ($r = -0.328$; $p = 0.158$).

Repulsive neural biases between trials away from sensory history

To probe for neural between-trial biases, we used the same approach as for within-trial neural biases. We tested the LDA evidence derived during the stimulus encoding period for systematic deviations in likelihood estimations as a function of the angular distance between the current orientation and the probed orientation on the previous trial. If the behavioral between-trial bias reflects neural modulation during encoding of sensory features, we would expect to see an increase in likelihood for orientations presented in the previous trial, in line with the attractive performance bias. We trained the classifier on data following presentation of both the first and second grating orientation combined and included all cue types and number of items presented. Informed by our behavioral analyses on between-trial performance biases, for the test set we selected trials where the previous probe angle had a relative difference of 0–60° (derived from significant angular differences; Fig. 3A, B) positive or negative from the presented grating orientation. The results were qualitatively the same and remained significant when other angular ranges were selected.

Contrary to our expectations, classifier evidence was significantly shifted away from the target orientation on the previous

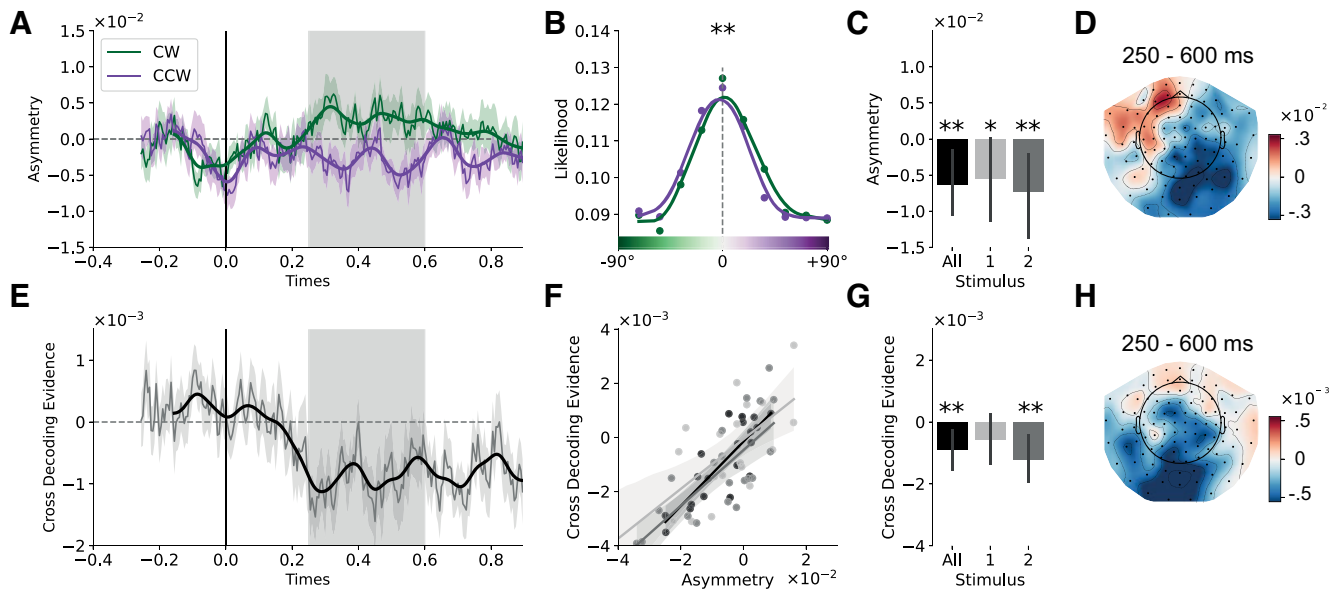


Figure 6. Bias on encoding imposed by previous trial. **A**, Asymmetry index for trials with a CW or CCW angular distance (0–60°) between the cued item on the previous trial and the presented grating (collapsed across the first and second grating). Asymmetry index time course is shown relative to the onset of grating presentation. A one-dimensional Gaussian filter with a kernel of 15 ms was applied to the data for visualization purposes. The gray shaded area indicates which time points (250–600 ms) are used for analyses described in the text and in panels **B–D** and **F–H**. **B**, Average LDA orientation-likelihoods for trials with CW (green) and CCW (purple) angular distances with respect to the previous trial. Evidence is lower for CW orientation bins when in the previous trial a CW orientation was cued, and vice versa for CCW orientations. The same interpolation was applied as described in Figure 5. **C**, Average bias relative to the previously cued orientation across all stimuli and for the first and second orientation separately. **D**, MEG topography based on searchlight results, illustrating in which sensors the evidence shift from panels **A–C** is most prominent. **E**, Cross-decoding evidence for the previously cued orientation (classifier trained on current orientation), locked to grating presentation. Same conventions as in panel **A**. **F**, Correlations between asymmetry index panel **A** and cross-decoding evidence from panel **E** for grating orientation processing of all stimuli combined (black), first grating (light gray) and second grating (dark gray). A least-squares linear regression model was applied for each of the conditions, with shaded 95% confidence intervals estimated using bootstrapping. **G**, Cross-decoding evidence from panel **D** summarised for all stimuli and separately for the first and second grating. **H**, MEG topography based on searchlight results, showing cross-decoding evidence for each sensor. * $p < 0.05$, ** $p < 0.01$. Plotted with 95% confidence intervals.

trial (250–600 ms postgrating onset; $Z = -3.228$, $p = 0.001$; Fig. 6A–D) rather than mirroring the attractive behavioral bias. In practice, this would mean that if the cued orientation on the previous trial was CW, classifier evidence for CCW bins increased, and vice versa. The repulsive bias away from the target on the previous trial was significant for the first ($Z = -2.020$, $p = 0.043$) and for the second grating ($Z = -2.690$, $p = 0.007$) in the current trial when considered separately (Fig. 6C). The between-trial repulsive bias during stimulus-two processing was present if no orientation was presented in the first interval ($Z = -2.298$, $p = 0.021$) but not when the first grating was also presented ($Z = -1.714$, $p = 0.087$). Interestingly, topographies in Figure 6D,G were highly similar ($r = 0.563$; $p < 0.001$), and both topographies correlated negatively with the stimulus-decoding topography [$r = 0.558$, $p < 0.001$ (Fig. 6D); $r = 0.763$, $p < 0.001$ (Fig. 6G)].

Next, we tested whether this repulsive neural bias was affected by the task relevance of the grating and by the cueing condition of the previous trial. We quantified this bias for task-relevant and task-irrelevant grating orientations in the previous trial. Only trials where the previous trial contained two items were included in this analysis. When averaging the neural bias over 250–600 ms, the task-relevant orientation showed a repulsive neural bias ($Z = -2.565$, $p = 0.010$), but we found no neural bias for task-irrelevant orientations ($Z = -0.174$, $p = 0.861$), though this difference did not reach significance ($t_{(19)} = 1.80$, $p = 0.088$). Yet, cluster-based permutation testing indicated a significant cluster indicating that task-relevant orientations exerted a significantly stronger repulsive bias than task-irrelevant orientations (500–550 ms; $p = 0.039$).

Cross-decoding evidence for previously presented stimuli

If information about the previously presented orientation is still partially present in the visual system during and after the presentation of the current grating orientation, it could interact with encoding, possibly leading to the observed repulsive bias. One possibility is that the lingering representation is in an orthogonal representational format, which is different from sensory coding of features (Libby and Buschman, 2021). In this case there would be little to no overlap between the activation pattern elicited during sensory input and the pattern related to the lingering representation of that past grating orientation. A classifier trained to separate perceptual information would therefore not cross-generalize if tested on the memory code. Alternatively, the lingering code could be present in a stable representation that shares similarity with the representation of incoming sensory information. If this were the case, a classifier, trained on the data from the current grating orientation, would cross-generalize and identify information about the past grating (or suppression of information expressed in negative evidence).

We adapted cross-decoding to detect lingering orientation-selective activity from the previous trial (also see Wan et al., 2020). After training the classifier for the presented orientation and testing for evidence of the previous trial's target orientation, we observed significant negative classifier evidence in the period of 250–600 after grating onset (concatenating over grating one and grating two: $t_{(19)} = -3.078$, $p = 0.006$; grating one alone, $t_{(19)} = -1.376$, $p = 0.185$; grating two, $t_{(19)} = -2.951$, $p = 0.008$; Fig. 6F).

Negative classifier evidence indicates that, while information is still present about the previous orientation, orientation-selective patterns may be sign-reversed relative to stimulus encoding.

This suppression of evidence for the previous trial's orientation could have been the cause of the apparent repulsive bias in the decoding of the current trial's orientation. If this was the case, we would expect the two measures to be positively correlated: stronger suppression of the previous orientation (negative classifier evidence) should lead to a more negative bias for the current orientation. We tested this using Pearson correlations across participants and found a significant correlation when assessing all grating presentations together ($r = 0.832$, $p < 0.001$; grating one alone, $r = 0.685$, $p < 0.001$; grating two alone, $r = 0.751$, $p < 0.001$; Fig. 6E).

No correlation was observed between the attractive behavioral performance bias and the magnitude of the negative shift in the neural data ($r = 0.186$, $p = 0.432$) nor with the magnitude of negative decoding ($r = 0.144$, $p = 0.544$).

Discussion

The present study investigated biases from previously perceived and memorized information on neural coding and behavioral responses. We observed both neural and behavioral biases that demonstrate interactions between past and present sensory processing. Orientations presented on the same trial exerted repulsive performance biases on currently perceived orientations. In contrast, task-relevant orientations on the previous trial exerted an opposite, attractive performance bias, altogether providing evidence for two counteracting biasing processes acting in tandem. Interestingly, multivariate decoding of neural data indicated that both stimuli from the same and previous trial generate a repulsive neural bias, suggesting that the two types of performance biases may arise from modulations acting on different stages of stimulus processing. While repulsive biases could reflect mechanisms akin to visual adaptation that promote visual discriminability, the attractive performance bias observed across trials was not observed at the level of the sensory representation. Since a neural repulsive bias occurred instead, we speculate that postperceptual modulatory mechanisms may override any early repulsive sensory modulation and lead to attractive performance biases.

The present behavioral results provide evidence of two types of performance biases: a repulsive within-trial and an attractive between-trial performance bias (cf. Bae and Luck, 2020; Fischer et al., 2020). Both biases were modulated by the task relevance of the inducing stimulus. First, we identified a repulsive performance bias away from the first orientation in a trial, but no significant retrospective repulsive bias from the second orientation, although the second grating was presented closer in time to the probe. We speculate that this may have happened because the relevance cue appeared together with this second stimulus, indicating its irrelevance in probe-first trials. Similarly, participants' responses in the current trial were biased only toward task-relevant orientations in the previous trial. At the neural level, biases were also dependent on task relevance. A repulsive neural bias was only present when the presented orientation was cued as task-relevant and therefore encoded into working memory. By contrast, task-irrelevant gratings that were not encoded into working memory could be decoded with similar precision but did not exhibit a significant bias. Gratings from previous trials that were associated with an attractive performance bias also led to repulsive neural biases during working-memory encoding. Only task-relevant orientations led to a repulsive bias.

Adaptation could partly explain the repulsive neural bias observed here. Visual adaptation has been proposed as the cause

of repulsive neural biases (Jazayeri and Movshon, 2006, 2007; Kohn, 2007; Stocker and Simoncelli, 2007; Webster, 2015). Since adaptation reduces firing in recently active neurons (Clifford et al., 2000; Wainwright, 1999), it is an efficient use of finite neural resources when the environment is autocorrelated (Stocker and Simoncelli, 2007; Webster, 2015) because neurons can code for a larger range of stimuli when their responses are not saturated. Curiously, we observed a repulsive neural bias on report second trials only when the presented orientation was task-relevant and encoded into working memory. The interaction with task relevance suggests that stimulus processing is only biased when it is primed for use in upcoming behavior.

This task-dependent modulation was unlikely the result of reduced processing of task-irrelevant stimuli, as overall orientation decoding was not affected by task relevance. In line with recent studies, it is possible that a context-sensitive repulsive bias, possibly occurring at a postperceptual stage (Zamboni et al., 2016; Fritsche and de Lange, 2019), exists alongside an early perceptual bias based on visual adaptation (Fritsche et al., 2017). The task-dependency of the neural bias could be the basis of recently observed repulsive biases in behavior (Bae and Luck, 2017; Czoschke et al., 2019, 2020; Chunharas et al., 2022), which may help individuate concurrently maintained stimuli (Wei et al., 2012). Under this explanation, only attended and encoded features would be subject to interactions with previous features.

The bias imposed by the orientation from the previous trial was attractive in behavior but repulsive in the neural data at early processing stages. This contrast is ostensibly at odds with previous behavioral studies that have assigned an early perceptual origin to attractive between-trial biases (Fischer and Whitney, 2014; Cicchini et al., 2017, 2018). There is still little direct neural evidence confirming this. EEG studies have shown that previous trial information can be decoded during the encoding phase of the current trial (Bae and Luck, 2019) or immediately before the current trial (Barbosa et al., 2020), and visually evoked neural responses in numerosity judgment tasks are modulated by stimulus history (Fornaciai and Park, 2018, 2020). However, the mere presence of prior stimulus information does not imply an attractive bias on the current stimulus. One study did observe an attractive behavioral bias and a neural bias in early visual areas using fMRI (St. John-Saaltink et al., 2016), but the study used only two stimulus orientations (45° and 135°) with an offset too large to produce a reliable behavioral bias, meaning that our findings may reflect a different biasing phenomenon.

Contradicting the early sensory origin of serial attractive bias, a recent fMRI study (Sheehan and Serences, 2022) found evidence consistent with the present results. Sheehan and Serences observed repulsive neural biases relative to the orientation on the previous trial across visual cortex, despite an attractive behavioral bias, and found that models incorporating early visual adaptation and a postperceptual origin of attractive biases could explain both effects. These results are broadly in line with our findings. We found that the repulsive neural bias of the classified orientation emerged relatively early in the trial, consistent with early visual cortex as the source of the bias. Our primary analysis approach did not allow us to infer anatomic sources of the bias (since we included data from all sensors as features in the classifier and their anatomic interpretability was further reduced by the dimensionality reduction step). We therefore conducted a searchlight analysis to identify which sensors contributed most to stimulus classification and to the repulsive bias. These spatially resolved results were in line with the visual cortical site of the repulsive bias observed by Sheehan and

Serences. Together, the observations make a case that prior stimuli lead to repulsion at the encoding stage and that attractive performance biases may arise from a different type of bias that did not systematically affect orientation decoding. Our findings further indicate that the link between neural adaptation and behavior can be context dependent, since repulsive neural biases could lead to both repulsive (within-trial) and attractive (between-trial) behavioral biases. Therefore, future models linking visual adaptation to behavior may need to incorporate context-dependence. Additionally, the high temporal resolution of MEG allowed us to show that neural biases arise within 500 ms of stimulus onset, have a posterior origin, and that they occur simultaneously relative to multiple prior stimuli (from the same trial and the previous trial).

While Sheehan and colleagues (Sheehan and Serences, 2022) argued that past stimuli were stored in a nonsensory code, they did not directly examine whether the representation of past stimuli occurred in a shared neural subspace with the representation of current stimuli. Here, we addressed this issue by showing that neural suppression of recently active neural populations could account for the observed repulsive bias. We tested this using cross-decoding analyses, training a classifier on the presented orientation and predicting previous orientations. Consistent with the suppression of recent stimulus-specific activity, cross-decoding yielded below-chance decoding of the previous orientation. In turn, this may have shifted the neural tuning curve for the current orientation away from the previous orientation, generating a repulsive bias. This relationship was confirmed by the robust correlation between cross-decoding and repulsive bias magnitude.

Another recent study observed a repulsive neural bias in single-unit recordings from frontal eye fields (FEF) paired with an attractive behavioral bias in a delayed-saccade task (Papadimitriou et al., 2017). The authors suggested that lingering attention to the previous target location could warp the representation of current target locations (Zirnsak et al., 2014). Since we observed neural biases primarily in posterior sensors, our results are more in line with a sensory origin of the repulsive bias, but this may interact with attentional biases originating in frontal cortex (Moore and Armstrong, 2003; Taylor et al., 2007). Attentional modulation could be one explanation for the context-dependency of biases observed here.

The current study, along with the study by Sheehan and Serences (2022) provides evidence against an early sensory origin of the attractive serial bias, but it does not provide direct evidence for how it arises. We speculate that the attractive bias does not manifest itself in orientation decoding or is not strong enough to overcome the repulsive bias during encoding. The attractive bias could originate postdecoding but since we were not able to identify stimulus-specific activity patterns later in the maintenance delay, we cannot say how this occurs. One possibility is that serial biases in behavior are not caused by biased neural representations at all. Akrami et al. (2018; also Boboeva et al., 2023) have suggested that prior stimulus information is stored in parietal neural activity, separately from the maintenance of the current memory stimulus. Its influence on behavior only emerges when a response is made that draws on both neural signals, thus creating an attractive bias. Bayesian accounts of serial dependence (Fritsche et al., 2020), while conceptually distinct, could equally rely on the separate maintenance and decision-stage integration of the prior and current stimuli. Such models could explain why there is so little positive evidence for an attractive neural bias at any processing stage.

Altogether, we demonstrate a consistent repulsive shift in neural evidence during working-memory encoding. Our results imply

that perceptual adaptation, along with context-sensitive factors, contributes to feature-selective downweighting to exert a repulsive bias away from recent stimulus features. Interestingly, no evidence of an attractive neural bias acting directly on sensory aspects of encoding was observed. Neural data thereby provide indirect evidence for the postperceptual account of attractive between-trial biases, rather than modulating encoding stages (Bliss et al., 2017; Fritsche et al., 2017; Pascucci et al., 2019; Bae and Luck, 2020; Kim et al., 2020). We speculate that the attractive between-trial bias instead arises through postperceptual processing stages involving memory (Bliss et al., 2017; Fritsche et al., 2017), perceptual decision-making, or motor planning (Boettcher et al., 2021; de Azevedo Neto and Bartels, 2021). The source of the attractive between-trial bias, whatever its neural mechanism, may be strong enough to override the repulsive bias during the perceptual/encoding stage. Together, these co-existing biases may help guide efficient coding for nuanced perceptual discriminations and visual stability across our environment.

References

- Akrami A, Kopec CD, Diamond ME, Brody CD (2018) Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature* 554:368–372.
- Bae GY, Luck SJ (2017) Interactions between visual working memory representations. *Atten Percept Psychophys* 79:2376–2395.
- Bae GY, Luck SJ (2019) Reactivation of previous experiences in a working memory task. *Psychol Sci* 30:587–595.
- Bae GY, Luck SJ (2020) Serial dependence in vision: merely encoding the previous-trial target is not enough. *Psychon Bull Rev* 27:293–300.
- Barbosa J, Stein H, Martinez RL, Galan-Gadea A, Li S, Dalmau J, Adam KCS, Valls-Solé J, Constantinidis C, Compte A (2020) Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nat Neurosci* 23:1016–1024.
- Bays PM, Catalao RFG, Husain M (2009) The precision of visual working memory is set by allocation of a shared resource. *J Vis* 9:7.1–11.
- Bliss DP, Sun JJ, D'Esposito M (2017) Serial dependence is absent at the time of perception but increases in visual working memory. *Sci Rep* 7:14739.
- Boboeva V, Pezzotta A, Clopath C, Akrami A (2023) From recency to central tendency biases in working memory: a unifying network model. *bioRxiv* 491352. <https://doi.org/10.1101/2022.05.16.491352>.
- Boettcher SE, Gresch D, Nobre AC, van Ede F (2021) Output planning at the input stage in visual working memory. *Sci Adv* 7:eabe8212.
- Born RT, Tootell RBH (1992) Middle temporal visual area. *Nature* 357:497–499.
- Brainard DH (1997) The psychophysics toolbox. *Spatial Vis* 10:433–436.
- Chunharas C, Rademaker RL, Brady TF, Serences JT (2022) An adaptive perspective on visual working memory distortions. *J Exp Psychol Gen* 151:2300–2323.
- Cicchini GM, Anobile G, Burr DC (2014) Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proc Natl Acad Sci U S A* 111:7867–7872.
- Cicchini GM, Mikellidou K, Burr D (2017) Serial dependencies act directly on perception. *J Vis* 17:6.
- Cicchini GM, Mikellidou K, Burr DC (2018) The functional role of serial dependence. *Proc R Soc B* 285:20181722.
- Clifford CW, Wenderoth P, Spehar B (2000) A functional angle on some after-effects in cortical vision. *Proc Biol Sci* 267:1705–1710.
- Czochke S, Fischer C, Beitzner J, Kaiser J, Bledowski C (2019) Two types of serial dependence in visual working memory. *Br J Psychol* 110:256–267.
- Czochke S, Peters B, Rahm B, Kaiser J, Bledowski C (2020) Visual objects interact differently during encoding and memory maintenance. *Atten Percept Psychophys* 82:1241–1257.
- de Azevedo Neto RM, Bartels A (2021) Disrupting short-term memory maintenance in premotor cortex affects serial dependence in visuomotor integration. *J Neurosci* 41:9392–9402.
- De Gardelle V, Kouider S, Sackur J (2010) An oblique illusion modulated by visibility: non-monotonic sensory integration in orientation processing. *J Vis* 10:6.

- Dong D, Atick J (1995) Statistics of natural time-varying images. *Netw Comput Neural Syst* 6:345–358.
- Fischer C, Czoschke S, Peters B, Rahm B, Kaiser J, Bledowski C (2020) Context information supports serial dependence of multiple visual objects across memory episodes. *Nat Commun* 11:1932.
- Fischer J, Whitney D (2014) Serial dependence in visual perception. *Nat Neurosci* 17:738–743.
- Fornaciai M, Park J (2018) Attractive serial dependence in the absence of an explicit task. *Psychol Sci* 29:437–446.
- Fornaciai M, Park J (2020) Neural dynamics of serial dependence in numerosity perception. *J Cogn Neurosci* 32:141–154.
- Fritsche M, de Lange FP (2019) Reference repulsion is not a perceptual illusion. *Cognition* 184:107–118.
- Fritsche M, Mostert P, de Lange FP (2017) Opposite effects of recent history on perception and decision. *Curr Biol* 27:590–595.
- Fritsche M, Spaak E, de Lange FP (2020) A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. *Elife* 9:e55389.
- Gekas N, McDermott KC, Mamassian P (2019) Disambiguating serial effects of multiple timescales. *J Vis* 19:24.
- Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, Hämäläinen M (2013) MEG and EEG data analysis with MNE-Python. *Front Neurosci* 7:267.
- Hajonides JE, Nobre AC, van Ede F, Stokes MG (2021) Decoding visual colour from scalp electroencephalography measurements. *Neuroimage* 237:118030.
- Huang J, Sekuler R (2010) Distortions in recall from visual memory: two classes of attractors at work. *J Vis* 10:24.1–27.
- Huang L (2020) Distinguishing target biases and strategic guesses in visual working memory. *Atten Percept Psychophys* 82:1258–1270.
- JASP Team (2020) JASP (version 0.14.1) [Computer software].
- Jazayeri M, Movshon JA (2006) Optimal representation of sensory information by neural populations. *Nat Neurosci* 9:690–696.
- Jazayeri M, Movshon JA (2007) A new perceptual illusion reveals mechanisms of sensory decoding. *Nature* 446:912–915.
- Kim S, Burr D, Cicchini GM, Alais D (2020) Serial dependence in perception requires conscious awareness. *Curr Biol* 30:R257–R258.
- Kiyonaga A, Scimeca JM, Bliss DP, Whitney D (2017) Serial dependence across perception, attention, and memory. *Trends Cogn Sci* 21:493–497.
- Kohn A (2007) Visual adaptation: physiology, mechanisms, and functional benefits. *J Neurophysiol* 97:3155–3164.
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868.
- Libby A, Buschman TJ (2021) Rotational dynamics reduce interference between sensory and memory representations. *Nat Neurosci* 24:715–726.
- Moore T, Armstrong KM (2003) Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421:370–373.
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
- Papadimitriou C, White RL, Snyder LH (2017) Ghosts in the machine II: neural correlates of memory interference from the previous trial. *Cereb Cortex* 27:2513–2527.
- Pascucci D, Mancuso G, Santandrea E, Libera CD, Plomp G, Chelazzi L (2019) Laws of concatenated perception: vision goes for novelty, decisions for perseverance. *PLoS Biol* 17:e3000144.
- Rosner B (1983) Percentage outlier points for generalized ESD many-procedure. *Technometrics* 25:165–172.
- Schneegans S, Bays PM (2017) Neural architecture for feature binding in visual working memory. *J Neurosci* 37:3913–3925.
- Sheehan TC, Serences JT (2022) Attractive serial dependence overcomes repulsive neuronal adaptation. *PLoS biology* 20:e3001711.
- Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
- St. John-Saaltink E, Kok P, Lau HC, De Lange FP (2016) Serial dependence in perceptual decisions is reflected in activity patterns in primary visual cortex. *J Neurosci* 36:6186–6192.
- Stocker AA, Simoncelli EP (2007) A Bayesian model of conditioned perception. *Adv Neural Inf Process Syst* 2007:1409–1416.
- Störmer VS, Alvarez GA (2014) Feature-based attention elicits surround suppression in feature space. *Curr Biol* 24:1985–1988.
- Taulu S, Kajola M, Simola J (2004) Suppression of interference and artifacts by the signal space separation method. *Brain Topogr* 16:269–275.
- Taylor PC, Nobre AC, Rushworth MF (2007) FEF TMS affects visual cortical activity. *Cereb Cortex* 17:391–399.
- Tomassini A, Morgan MJ, Solomon JA (2010) Orientation uncertainty reduces perceived obliquity. *Vision Res* 50:541–547.
- van Bergen RS, Jehee JF (2019) Probabilistic representation in human visual cortex reflects uncertainty in serial decisions. *J Neurosci* 39:8164–8176.
- Virtanen P, et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 17:261–272.
- Wainwright MJ (1999) Visual adaptation as optimal information transmission. *Vision Res* 39:3960–3974.
- Wan Q, Cai Y, Samaha J, Postle BR (2020) Tracking stimulus representation across a 2-back visual working memory task: tracking 2-back representation. *R Soc Open Sci* 7:190228.
- Webster MA (2015) Visual adaptation. *Annu Rev Vis Sci* 1:547–567.
- Wei Z, Wang XJ, Wang DH (2012) From distributed resources to limited slots in multiple-item working memory: a spiking network model with normalization. *J Neurosci* 32:11228–11240.
- Wolff MJ, Jochim J, Akyürek EG, Stokes MG (2017) Dynamic hidden states underlying working-memory-guided behavior. *Nat Neurosci* 20:864–871.
- Wolff MJ, Jochim J, Akyürek EG, Buschman TJ, Stokes MG (2020) Drifting codes within a stable coding scheme for working memory. *PLoS Biol* 18:e3000625.
- Zamboni E, Ledgeway T, McGraw PV, Schluppeck D (2016) Do perceptual biases emerge early or late in visual processing? Decision-biases in motion perception. *Proc Biol Sci* 283:20160263.
- Zirnsak M, Steinmetz NA, Noudoost B, Xu KZ, Moore T (2014) Visual space is compressed in prefrontal cortex before eye movements. *Nature* 507:504–507.